Application of Leftover Hash Lemma to Public Key Encryptions Based on Conjugacy Search Problem (New contact points of algebraic systems, logics, languages, and computer sciences)

AUTHOR(S):

山村, 明弘

数理解析研究所講究録
第1964巻 2015年 40-45

Application of Leftover Hash Lemma to Public Key Encryptions Based on Conjugacy Search Problem *

Akihiro Yamamura
Akita University
1-1, Tegata Gakuen-machi, Akita 010-8502 Japan
yamamura@ie.akita-u.ac.jp

Abstract

We show flaws in a one of public key encryptions based on conjugacy search problem which has fatal flaw and apply the leftover hash lemma to remedy.

1 Flaws CSP-ElG

Three public key encryptions, CSP-ElG, CSP-hElG and CSP-CS schemes, are proposed by L.Wang, L.Wang, Z.Cao, E.Okamoto and J.Shao in Inscrypt 2010 [7]. Each scheme is claimed to have certain provable security. It is reported a fatal flaw in CSP-ElG and indicated that there are more erros in the design of the other two schemes in [8]. In this paper, we review the results obtained in [8] and consider the future research problems.

Let M be a (not necessarily commutative) monoid. We denote the set of invertible elements x of M by $G(M)$. The conjugacy search problem is to find an element $g \in G(M)$ such that $f = gdg^{-1}$ for given $d, f \in M$ provided that such an element g exists.

Suppose that $d \in M$ and $g \in G(M)$ and the order of g is n. If the order of g is infinite, then n is specified to be a large enough. The CSP-DDH problem is a decisional problem to decide whether or not $f = g^{a+b}dg^{-(a+b)}$ for given $d \in M$, $g \in G(M)$, $g^a dg^{-a}$, $g^b dg^{-b}$ and $f = g^c dg^{-c}$, where a and b and are randomly chosen from $\{1, \ldots, n\}$ and either c is randomly chosen from $\{1, \ldots, n\}$ or $c = a + b$ with probability $\frac{1}{2}$. We say that the CSP-DDH assumption holds for M if there is no efficient algorithm to answer correctly

*This work was supported by JSPS KAKENHI Grant Number 15K00004

to a CSP-DDH problem instance with probability non-negligibly larger than $\frac{1}{2}$.

The CSP-ElG scheme is defined as follows. Let $K = \{g^a dg^{-a} \mid 1 \leq a \leq \mathrm{Ord}(g)\}$, where $\mathrm{Ord}(g)$ stands for the order of the element g. Suppose $H : K \to P$ is a cryptographic hash function. Let P be the message space $\{0,1\}^k$, C the ciphertext space $K \times P$. Alice picks a $(1 \leq a \leq \mathrm{Ord}(g))$ and publicizes $g^a dg^{-a}$. Bob picks b $(1 \leq b \leq \mathrm{Ord}(g))$ and encrypts a message $m \in P$ by

$$c = (g^b dg^{-b}, m \oplus H(g^b(g^a dg^{-a})g^{-b})).$$

Receiving the ciphertext $c = (c_1, c_2)$, Alice decrypts it by $m = c_2 \oplus H(g^a c_1 g^{-a})$.

Theorem 1 of [7] claims that the CSP-ElG scheme is indistinguishable against chosen plaintext attacks in the standard model. On the other hand, we must not assume an random oracle in the standard model, and so we may not assume H is a random oracle. We shall see that if H is a random oracle, the scheme is indistinguishable against chosen plaintext attacks and a random oracle is vital in the CSP-ElG scheme and this disproves Theorem 1 of [7].

We choose two messages m_1 and m_2 from P. One of them is chosen by coin toss and it is encrypted as c then we are asked to decide whether c is a ciphertext of m_1 or m_2. First, we define a cryptographic hash function H to be

$$H(m) = \text{SHA-1}(m)|0. \tag{1.1}$$

The value of H is the concatenation of the value of SHA-1(m) and a bit 0. Then H is a cryptographic hash function of hash size 161 bits and satisfies collision resistance, pre-image and second pre-image resistance, while it is not a random oracle because the last bit is always 0 and so the hash value is not random.

Let $P = \{0,1\}^{161}$. Take m_1 as any message with the last bit is 1, and m_2 as any message with the last bit is 0. Then the ciphertext of m_1 is given by

$$c = (g^b dg^{-b}, m_1 \oplus H(g^b(g^a dg^{-a})g^{-b})).$$

The last bit of the second entry is 1 since the last bit of $H(g^b(g^a dg^{-a})g^{-b})$ is 0. The ciphertext of m_2 is

$$c = (g^b dg^{-b}, m_2 \oplus H(g^b(g^a dg^{-a})g^{-b})).$$

Similarly the last bit of the second entry is 0 since the last bit of $H(g^b(g^a dg^{-a})g^{-b})$ is 0. Therefore an attacker can always distinguish the ciphertexts of m_1 and m_2 with probability 1. This shows that the CSP-ElG scheme is not indistinguishable in the standard model and disproves Theorem 1 of [7].

2 Gennaro, Krawczyk and Rabin's Method

We recall Gennaro, Krawczyk and Rabin's method to obtain a uniform distribution over the set $\{0,1\}^s$ of the fixed length bit strings from DH transforms over non-DDH groups for preparation for our fixing the CSP-ElG scheme. The ElGamal encryption is indistinguishable against chosen plaintext attacks provided a generator g is chosen adequately and the base group enjoys the DDH assumption. However, g may be chosen inadequately and its order may be insufficient in length in real-life systems. For example, SSH and IPSec standards instantiate groups in which the DDH assumption does not necessarily hold. Even in such a case, the ElGamal scheme still enjoys provable security under the so-called t-DDH assumption introduced in [3].

We recall necessary terminology. Let $\mathcal{X}$ and $\mathcal{Y}$ be random variables with support contained in $\{0,1\}^n$. The *statistical distance* between $\mathcal{X}$ and $\mathcal{Y}$ is

$$\mathrm{dist}(\mathcal{X},\mathcal{Y}) = \frac{1}{2}\sum_{x\in\{0,1\}^n} |\mathrm{Prob}(\mathcal{X}=x) - \mathrm{Prob}(\mathcal{Y}=x)|.$$

Now suppose $\mathcal{X}_n$ and $\mathcal{Y}_n$ are probability ensembles. Let $\mathcal{D} = \{D_n\}$ be a family of circuits. Then $\mathcal{X}_n$ and $\mathcal{Y}_n$ are called *computationally indistinguishable* (by non-uniform distinguishers) if for every polynomial-size distinguisher family $\mathcal{D}$, for every polynomial $P(\cdot)$ and for sufficiently large n we have

$$|\mathrm{Prob}_{x\in\mathcal{X}_n}(D_n(x)=1) - \mathrm{Prob}_{y\in\mathcal{Y}_n}(D_n(y)=1)| \le \frac{1}{P(n)}.$$

Let $\mathcal{X}_n$ be a probability ensemble over A_n. The *min-entropy* of $\mathcal{X}_n$ is defined to be

$$\mathsf{min\text{–}ent}(\mathcal{X}_n) = \mathrm{min}_{x\in A_n:\mathrm{Prob}_{x\in\mathcal{X}_n}(x)\neq 0}(-\log(\mathrm{Prob}_{x\in\mathcal{X}_n}(x))).$$

Let $\mathcal{G} = \{G_n\}$ be a family of cyclic groups. We say that $t(n)$-*DDH assumption* holds over $\mathcal{G}$ if for all n there exists a family of probability distributions $\mathcal{X}_n(x^a, x^b)$ such that

1. $\mathsf{min\text{-}ent}(\mathcal{X}_n(x^a,x^b)) \ge t(n)$
2. The probability ensemble

$$\mathcal{DH}_n = \{(x^a, x^b, x^{ab} \mid a,b \in_U \{1,\ldots,\mathrm{Ord}(G_n)\}\}$$

is computationally indistinguishable from the ensemble

$$\mathcal{R}^*_n = \{(x^a, x^b, C \mid a,b \in_U \{1,\ldots,\mathrm{Ord}(G_n)\} \text{ and } C \in_{\mathcal{X}_n(x^a,x^b,x^{ab})} G_n\},$$

where $\mathrm{Ord}(G_n)$ stands for the order of the group G_n.

The notation $x \in_{\mathcal{D}} A$ is to be read as x is chosen from A according to the distribution $\mathcal{D}$, and $x \in_U S$ means choosing x uniformly from the set S. The probability distributions $\mathcal{X}_n(x^a, x^b)$ may be different for each triple x, x^a, x^b. Intuitive meaning of the assumption is that a DH output x^{ab} has some degree of unpredictability.

2.1 Universal hashing

Universal hashing was introduced by Carter and Wegman [2] in 1979 and has been a basic technique in many areas of information security. It realizes pseudorandom generators, privacy amplification and derandomization. Universal hashing is formed by orthogonal arrays and error correcting codes, and so on [6]. Leftover hash lemma was given by Impagliazzo, Levin and Luby [5] and has many applications together with universal hashing. See also [1]. We use universal hashing to correct one of public key encryptions based on conjugacy search problem proposed in Inscrypt 2010 [7].

Suppose $h : \{0,1\}^n \times \{0,1\}^{l(n)} \to \{0,1\}^{m(n)}$ is a function. For each fixed $Y \in \{0,1\}^{l(n)}$ we have a function $h_Y(\,\cdot\,) = h(\,\cdot\,, Y)$ that maps n bits to $m(n)$ bits. Then h is called a *(pairwise independent) universal hash function* if for all $x_1, x_2 \in \{0,1\}^n$ $(x_1 \neq x_2)$ and for all $a_1, a_2 \in \{0,1\}^{m(n)}$, we have

$$\mathrm{Prob}_{Y \in_U \{0,1\}^{l(n)}}(h_Y(x_1) = a_1 \text{ and } h_Y(x_2) = a_2) = \frac{1}{2^{2m(n)}}.$$

Leftover hash lemma is introduced and used to construct pseudorandom bit strings in [4] and used to smooth distributions in [3]. See also [1] for a recent development of the leftover hash lemma.

Lemma 2.1 (Leftover hash lemma [4]) *Let $\mathcal{X}_n$ be a probability ensemble such that* $\mathsf{min\text{–}ent}(\mathcal{X}_n) = m(n)$. *Let $e(n)$ be a positive integer valued parameter. Let $h : \{0,1\}^n \times \{0,1\}^{l(n)} \to \{0,1\}^{m(n)-2e(n)}$ be a universal hash function. Let $X \in_{\mathcal{X}_n} \{0,1\}^n$, $Y \in_U \{0,1\}^{l(n)}$ and $Z \in_U \{0,1\}^{m(n)-2e(n)}$. Then we have*

$$\mathrm{dist}(\langle h_Y(X), Y\rangle, \langle Z, Y\rangle) \leq \frac{1}{e(n)+1},$$

where $\langle X, Y\rangle$ stands for the concatenation of X and Y.

Using the leftover hash lemma, Gennaro et al. [3] show that if $\mathcal{G} = \{G_n\}_n$ is a group family in which the $t(n)$-DDH assumption holds and $h : \{0,1\}^{|G_n|} \times \{0,1\}^{l(n)} \to \{0,1\}^{t'(n)}$ is a universal hash function, where $t'(n) = t(n) - \omega(\log n)$, then the induced distribution of $h(g_n^{ab}, Y)$ for $a, b \in_U$

$\{1, 2, \ldots, \mathrm{Ord}(G_n)\}$ and $Y \in_U \{0,1\}^{l(n)}$ is computationally indistinguishable from the uniform distribution over $\{0,1\}^{t'(n)}$ even when h, g_n^a and g_n^b are given to the distinguisher. This implies that the ElGamal scheme using the hashed value $h(g_n^{ab}, Y)$ instead of g_n^{ab} to mask a plaintext is indistinguishable if the underlying group satisfies $t(n)$-DDH assumption. In this case the universal hash function is common knowledge between Alice and Bob and $Y(\in \{0,1\}^{l(n)})$ is a piece of a ciphertext.

3 Revised CSP-ElG scheme

In [8], CSP-ElG scheme is revised. Suppose $t(n)$-CSP-DDH assumption (one of a concrete instance of the $t(n)$-MA-DDH assumption) holds for M, $d \in M$, $g \in G(M)$ and $h : \{0,1\}^{|M_n|} \times \{0,1\}^{l(n)} \to \{0,1\}^{t'(n)}$ is a universal hash function. A public key is a pair $(g, g^a d g^{-a})$. A plaintext $P \in \{0,1\}^{t'(n)}$ is encrypted as

$$(Y, g^b d g^{-b}, P \oplus h(Y, g^{a+b} d g^{-(a+b)})),$$

where $Y \in \{0,1\}^{l(n)}$. In this case, the universal hash function h is publicized and $Y(\in_U \{0,1\}^{l(n)})$ is a piece of a ciphertext.

We have the following the theorem. The reader is referred to [8] for detail proof.

Theorem 3.1 *The revised CSP-ElG scheme is indistinguishable against chosen plaintext attacks in the standard model if the $t(n)$-CSP-DDH assumption holds.*

4 CSP-hElG and CSP-CS

In addition to CSP-ElG, CSP-hElG and CSP-CS are proposed in [7]. The authors claim the schemes have provable security. However, it is no longer trustworthy after CSP-ElG has a security flaw. The missing argument is that the authors in [7] did not tell random oracles from collision resistent hash functions. In the case of CSP-ElG, we use the universal hash functions to remedy. It is plausible to use the same method to remedy CSP-hElG and CSP-CS. On the other hand, the proof of the security for CSP-hElG and CSP-CS would be extremely harder than that of CSP-ElG and so we may need some more idea. It is also necessary to obtain more algebraic systems with computationally hard CSP problem. Furthermore, the study of universal hashing is also of significance.

References

[1] B.Barak, Y.Dodis, H.Krawczyk, O.Pereira, K.Pietrzak, F.-X.Standaert and Y.Yu, Leftover hash lemma, revisited, CRYPTO 2011, LNCS Vol. 6841, (2011) 1–20.

[2] L.Carter and M.N.Wegman, Universal classes of hash functions, J. Computer and System Sciences, 18 (2), (1979) 143–154.

[3] R.Gennaro, H.Krawczyk and T.Rabin, Secure hashed Diffie-Hellman over non-DDH groups, EUROCRYPT 2004, LNCS Vol. 3027, (2004) 361–381.

[4] J.Hastad, R.Impagliazzo, L.Levin and M.Luby, Construction of a pseudo-random generator from any one-way function, SIAM J. Computing, 28 (4), (1999) 1364–1396.

[5] R.Impagliazzo, L.Levin, and M.Luby, Pseudo-random generation from one-way functions, STOC (1989) 12–24.

[6] D.R.Stinson, On the connections between universal hashing, combinatorial designs and error-correcting codes, Congressus Numerantium, Vol. 114 (1996) 7-27.

[7] L.Wang, L.Wang, Z.Cao, E.Okamoto and J.Shao, New Constructions of Public-Key Encryption Schemes from Conjugacy Search Problems, INSCRYPT 2010, LNCS Vol. 6584, (2011) 1–17.

[8] A. Yamamura, Security Analysis of Public Key Encryptions Based on Conjugacy Search Problem, ICT-Eurasia 2014, LNCS Vol. 8407, (2014) 554–563.